On the Capacity of a Class of MIMO Cognitive Radios

Sriram Sridharan, *Student Member, IEEE* and Sriram Vishwanath, *Member, IEEE*

Abstract— **Cognitive radios have been studied recently as a means to utilize spectrum in a more efficient manner. This paper focuses on the fundamental limits of operation of a MIMO cognitive radio network with a single licensed user and a single cognitive user. The channel setting is equivalent to an interference channel with degraded message sets (with the cognitive user having access to the licensed user's message). An achievable region and an outer bound is derived for such a network setting. It is shown that under certain conditions, the achievable region is optimal for a portion of the capacity region that includes sum capacity.**

I. INTRODUCTION

The design of radios to be "cognitive" has been identified by the Federal Communications Commission (FCC) as the next big step in better radio resource utilization [2]. The term "cognitive" has many different connotations both in analysis and in practice, but with two underlying common themes: *intelligence* built into the radio architecture coupled with *adaptivity*.

Cognitive radios have been studied under different model settings. The first models studied cognitive radios as a spectrum sensing problem [3][4][5][6]. Under this setting, the cognitive radio opportunistically uses licensed spectrum when the licensed users are sensed to be absent in that band. Problems encountered in this setup are threefold :

1) Sensing must be highly accurate to guarantee non interference with the licensed radio.
2) Control and coordination between the cognitive transmitter receiver pair is required to ensure the same spectrum is used, and finally
3) There are no QoS guarantees for the cognitive transmitter receiver pair.

Other models with different side information at the cognitive users have been studied. In [7] and [8], the authors study frequency coding by the cognitive transmitter by assuming non causal knowledge of the frequency use of the primary transmitter.

In this paper, we study cognition from an information theoretic setting where we assume that the cognitive transmitter knows the message of the licensed transmitter apriori. Such a model is interesting for two reasons : 1) It provides an upper limit, or equivalently a benchmark on the performance of systems where the cognitive radio gains a partial understanding of the licensed transmitter and 2) It allows us to understand the ultimate limits on the cognitive transmitter by giving it maximum information and allowing it to change its transmission and coding strategy based on all the information available at the licensed user. In essence, it enlarges the possible schemes that can be implemented at the cognitive radio, and 3) It lends itself to information theoretic analysis, being a setting where such tools can be applied to determine the performance limits of the system. Many other configurations, including the interference channel setting when the cognitive transmitter does not know the message of the licensed transmitter are multi-decade long open problems.

The goal of this paper is to study the fundamental limits of performance of cognitive radios. Along the lines of [9], we consider the model depicted in Figure 1. In this setting, we have an interference channel [10][11][12][13], but with degraded message sets, where the transmitter with a single message is called "legacy," "primary" or "dumb" and the transmitter with both messages termed the "cognitive" transmitter. Prior work on this model for the single antenna case is in [9][14][15][16].

In this paper, we study the performance of the cognitive radio model under a multiple antenna (MIMO) setting. Both the licensed and cognitive transmitter and receiver may have multiple antennas. MIMO is fast becoming the most common feature of wireless systems due to its performance benefits. Thus, it is important to study the capacity of cognitive radios under a MIMO setting. There are some instances where the methods used in this paper bears similarities with the methods used for the SISO setting. However, most of the proofs and techniques used here are distinct and considerably more involved than those used in [16]. In the SISO setting, it is possible to analyze the model for specific magnitudes of channels. This is not possible for the MIMO setting. We list some of the crucial differences between the methods used in this paper and the methods that have been used under the SISO setting.

1) In [16], the authors obtain the outer bound using conditional entropy inequality. This method cannot be extended to the MIMO setting.
2) We obtain the outer bound through a series of channel

transformations. Although the channel transformations are similar in spirit to those in [15], the actual transformations used are significantly different both in nature and in the mathematical proofs that accompany them. In [15], the authors reduce the channel to a broadcast channel where the combined transmitters have individual power constraints and the cognitive receiver has the message of the licensed user provided to it by a genie. The capacity region for such a variation of broadcast channel is not known in general. The authors solve for the capacity region of the broadcast channel using aligned channel techniques. On the other hand, we reduce the MIMO cognitive channel to a broadcast channel with sum power constraint and whose capacity region is now known [17][18][19]. We then use optimization techniques to compare the achievable scheme with the outer bound.

A. Main Contributions

In this paper, our main contributions include:

1. We find an achievable region for the Gaussian MIMO cognitive channel (MCC) in a fashion analogous to [9][15][16].

2. We find an outer bound on the capacity region of the MCC.

3. We show that, under certain conditions (that depend on the channel parameters), the outer bound is tight for a portion of the capacity region boundary, including points corresponding to the sum-capacity of the channel. Combining the two above, we characterize the sum capacity of this channel and a portion of its entire capacity region under certain conditions.

B. Organization

The rest of the paper is organized as follows. We describe the notations and system model in Section II. The main results are presented in Section III. In Section IV, we present an achievable region for the Gaussian MIMO cognitive channel (MCC). An outer bound on the capacity region is shown in Section V. The optimality of the achievable region for a portion of the capacity region (under certain conditions) is shown in Section VI. Numerical results are provided in Section VII. We conclude in Section VIII.

II. SYSTEM MODEL AND NOTATION

Throughout the paper, we use boldface letters to denote vectors and matrices. $|\mathbf{A}|$ denotes the determinant of matrix $\mathbf{A}$, while $\mathrm{Tr}(\mathbf{A})$ denotes its trace. For any general matrix or vector $\mathbf{X}$, $\mathbf{X}^\dagger$ denotes its conjugate transpose. $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. $\mathbf{X}^n$ denotes the row vector $(X(1), X(2), \ldots, X(n))$, where $X(i), i = 1, 2, \ldots, n$ can be vectors or scalars. The notation $\mathbf{H} \succeq \mathbf{0}$ is used to denote that a square matrix $\mathbf{H}$ is positive semidefinite. Finally, if $\mathbf{S}$ is a

set, then $\mathrm{Cl}(\mathbf{S})$ and $\mathrm{Co}(\mathbf{S})$ denote the closure and convex hull of $\mathbf{S}$ respectively.

We consider a MIMO cognitive channel shown in Figure 1. Let $n_{p,t}$ and $n_{p,r}$ denote the number of transmitter and receiver antennas respectively for the licensed user. Similarly, $n_{c,t}$ and $n_{c,r}$ denotes the number of transmitter and receiver antennas for the cognitive user.
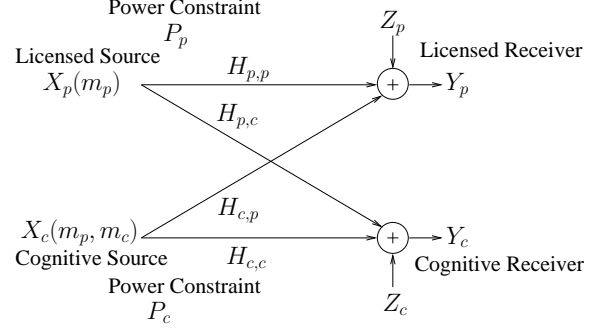

Fig. 1. MIMO Cognitive Radio System Model

The licensed user has message $m_p \in \{1, 2, \ldots, 2^{nR_p}\}$ intended for the licensed receiver. The cognitive user has message $m_c \in \{1, 2, \ldots, 2^{nR_c}\}$ intended for the cognitive receiver as well as the message m_p of the licensed user.

The primary user encodes the message m_p into $\mathbf{X_p}^n$. Here, $\mathbf{X_p}(i)$ is a $n_{p,t}$ length complex vector. The cognitive transmitter determines its codeword $\mathbf{X_c}^n$ as a function of both m_p and m_c. Note that the cognitive transmitter wishes to communicate both m_p (to the licensed receiver) and m_c (to the cognitive receiver). The channel gain matrices are given by $\mathbf{H_{p,p}}, \mathbf{H_{p,c}}, \mathbf{H_{c,p}}$ and $\mathbf{H_{c,c}}$, and are assumed to be static. It is assumed that the licensed receiver knows $\mathbf{H_{p,p}}, \mathbf{H_{c,p}}$, the licensed transmitter knows $\mathbf{H_{p,p}}$. It is also assumed that the cognitive transmitter knows $\mathbf{H_{c,p}}, \mathbf{H_{p,c}}, \mathbf{H_{c,c}}$ and the cognitive receiver knows $\mathbf{H_{p,c}}, \mathbf{H_{c,c}}$. The received vectors of the licensed and cognitive users are denoted by $\mathbf{Y_p}^n$ and $\mathbf{Y_c}^n$ respectively.

With the above model and notations, we can describe the system at time slot i by

$$\begin{aligned} \mathbf{Y_p}(i) &= \mathbf{H_{p,p}}\mathbf{X_p}(i) + \mathbf{H_{c,p}}\mathbf{X_c}(i) + \mathbf{Z_p}(i) \\ \mathbf{Y_c}(i) &= \mathbf{H_{p,c}}\mathbf{X_p}(i) + \mathbf{H_{c,c}}\mathbf{X_c}(i) + \mathbf{Z_c}(i). \end{aligned} \quad (1)$$

The additive noise at the primary and secondary receivers is denoted by $\mathbf{Z_p}^n$ and $\mathbf{Z_c}^n$ respectively. The noise vectors $\mathbf{Z_p}^n$ and $\mathbf{Z_c}^n$ are Gaussian and are assumed to be i.i.d. across symbol times and distributed according to $\mathcal{N}(0, \mathbf{I_{n_{p,r}}})$ and $\mathcal{N}(0, \mathbf{I_{n_{c,r}}})$ respectively. The correlation between $\mathbf{Z_p}^n$ and $\mathbf{Z_c}^n$ is assumed to be arbitrary. This correlation does not impact the capacity region of the system as the licensed and the cognitive decoders do not co-operate with each other. [1]

We denote the covariance of the codewords of the licensed and cognitive transmitters at time i by $\mathbf{\Sigma_p}(i)$ and $\mathbf{\Sigma_c}(i)$

---

[1] A proof of this can be obtained using steps almost exactly identical to those for the broadcast channel in [20, Exercise 15.10]

respectively. Then, the transmitters are constrained by the following transmit power constraints.

$$\sum_{i=1}^{n} \text{Tr}(\boldsymbol{\Sigma}_{\mathbf{p}}(i)) \leq nP_p \\ \sum_{i=1}^{n} \text{Tr}(\boldsymbol{\Sigma}_{\mathbf{c}}(i)) \leq nP_c. \tag{2}$$

A rate pair (R_p, R_c) is said to be achievable if

1) there exists a sequence of encoding functions for the licensed and cognitive users $E_p^n : \{1, \ldots, 2^{nR_p}\} \rightarrow \mathbf{X_p}^n$ and $E_c^n : \{1, \ldots, 2^{nR_p}\} \times \{1, \ldots, 2^{nR_c}\} \rightarrow \mathbf{X_c}^n$ such that the codewords satisfy the power constraints given by (2),
2) there exists decoding rules $D_p^n : \mathbf{Y_p}^n \rightarrow \{1, \ldots, 2^{nR_p}\}$ and $D_c^n : \mathbf{Y_c}^n \rightarrow \{1, \ldots, 2^{nR_c}\}$ such that the average probability of decoding error is arbitrarily small for suitably large values of n.

The capacity region of the Gaussian MIMO cognitive channel is the set of all achievable rate pairs (R_p, R_c) and is denoted by $\mathcal{C}_{MCC}$.

III. MAIN RESULTS

In this section, we describe the main results of the paper. Let $\mathbf{G} = [\mathbf{H_{p,p}} \quad \mathbf{H_{c,p}}]$. Let $\mathcal{R}_{ach}$ denote the set described by

$$\left\{ \begin{array}{l} \left((R_p, R_c), \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \mathbf{Q} \right) : \\ R_p \geq 0, R_c \geq 0, \boldsymbol{\Sigma_p} \succeq \mathbf{0}, \boldsymbol{\Sigma_{c,p}} \succeq \mathbf{0}, \boldsymbol{\Sigma_{c,c}} \succeq \mathbf{0} \\ R_p \leq \log \left| \mathbf{I} + \mathbf{G}\boldsymbol{\Sigma_{p,net}}\mathbf{G}^{\dagger} + \mathbf{H_{c,p}}\boldsymbol{\Sigma_{c,c}}\mathbf{H_{c,p}^{\dagger}} \right| \\ \qquad - \log \left| \mathbf{I} + \mathbf{H_{c,p}}\boldsymbol{\Sigma_{c,c}}\mathbf{H_{c,p}^{\dagger}} \right| \\ R_c \leq \log \left| \mathbf{I} + \mathbf{H_{c,c}}\boldsymbol{\Sigma_{c,c}}\mathbf{H_{c,c}^{\dagger}} \right| \\ \boldsymbol{\Sigma_{p,net}} = \left(\begin{array}{cc} \boldsymbol{\Sigma_p} & \mathbf{Q} \\ \mathbf{Q}^{\dagger} & \boldsymbol{\Sigma_{c,p}} \end{array} \right) \succeq \mathbf{0}, \\ \text{Tr}(\boldsymbol{\Sigma_p}) \leq P_p, \quad \text{Tr}(\boldsymbol{\Sigma_{c,p}} + \boldsymbol{\Sigma_{c,c}}) \leq P_c \end{array} \right\}. \tag{3}$$

In this setting, $\boldsymbol{\Sigma_{p,net}}$ is a $(n_{p,t} + n_{c,t}) \times (n_{p,t} + n_{c,t})$ covariance matrix while $\boldsymbol{\Sigma_{c,c}}$ is a $n_{c,t} \times n_{c,t}$ covariance matrix. $\boldsymbol{\Sigma_p}$ and $\boldsymbol{\Sigma_{c,p}}$ represent principal submatrices of $\boldsymbol{\Sigma_{p,net}}$ of dimensions $n_{p,t} \times n_{p,t}$ and $n_{c,t} \times n_{c,t}$ respectively. The covariances matrices $\boldsymbol{\Sigma_p}$, $\boldsymbol{\Sigma_{c,p}}$ and $\boldsymbol{\Sigma_{c,c}}$ determine the power constraints of the system.

Let $\mathcal{R}_{in}$ denote the closure of the convex hull of the set of rate pairs described by

$$\left\{ \begin{array}{l} (R_p, R_c) : \exists \ \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \mathbf{Q}, \text{ and} \\ \left((R_p, R_c), \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \mathbf{Q} \right) \in \mathcal{R}_{ach} \end{array} \right\}. \tag{4}$$

Theorem 3.1: The capacity region of the MCC, $\mathcal{C}_{MCC}$ satisfies

$$\mathcal{R}_{in} \subseteq \mathcal{C}_{MCC}. \tag{5}$$

The proof of the theorem is given in Section IV. The coding strategy is based on Costa's dirty paper coding [21][22].

We now describe an outer bound on the capacity region of the MIMO cognitive channel. Let $\alpha > 0$, $\mathbf{G}_\alpha = \left[\mathbf{H_{p,p}} \quad \frac{\mathbf{H_{c,p}}}{\sqrt{\alpha}} \right]$ and $\overline{\mathbf{K}} = \left[\begin{array}{cc} \mathbf{H_{p,p}} & \mathbf{H_{c,p}}/\sqrt{\alpha} \\ \mathbf{0} & \mathbf{H_{c,c}}/\sqrt{\alpha} \end{array} \right]$. Let $\boldsymbol{\Sigma_z}$ be a covariance matrix of dimensions $(n_{p,r} + n_{c,r}) \times (n_{p,r} + n_{c,r})$ and of the form

$$\boldsymbol{\Sigma_z} = \left[\begin{array}{cc} \mathbf{I}_{n_{p,r}} & \mathbf{Q_z} \\ \mathbf{Q_z}^{\dagger} & \mathbf{I}_{n_{c,r}} \end{array} \right]. \tag{6}$$

Here, $\mathbf{Q_z}$ is a $n_{p,r} \times n_{c,r}$ matrix that makes $\boldsymbol{\Sigma_z}$ positive semidefinite. Let $\mathcal{R}_{conv}^{\alpha, \boldsymbol{\Sigma_z}}$ denote the set described by

$$\left\{ \begin{array}{l} \left((R_p, R_c), \mathbf{Q_p}, \mathbf{Q_c} \right) : \\ R_p \geq 0, R_c \geq 0, \mathbf{Q_p} \succeq \mathbf{0}, \mathbf{Q_c} \succeq \mathbf{0} \\ R_p \leq \log \left| \mathbf{I} + \mathbf{G}_\alpha \mathbf{Q_p} \mathbf{G}_\alpha^{\dagger} + \mathbf{G}_\alpha \mathbf{Q_c} \mathbf{G}_\alpha^{\dagger} \right| \\ \qquad - \log \left| \mathbf{I} + \mathbf{G}_\alpha \mathbf{Q_c} \mathbf{G}_\alpha^{\dagger} \right| \\ R_c \leq \log \left| \boldsymbol{\Sigma_z} + \overline{\mathbf{K}} \mathbf{Q_c} \overline{\mathbf{K}}^{\dagger} \right| - \log |\boldsymbol{\Sigma_z}| \\ \text{Tr}(\mathbf{Q_p}) + \text{Tr}(\mathbf{Q_c}) \leq P_p + \alpha P_c \end{array} \right\}. \tag{7}$$

Let $\mathcal{R}_{out}^{\alpha, \boldsymbol{\Sigma_z}}$ denote the closure of the convex hull of the set of rate pairs described by

$$\left\{ (R_p, R_c) : \exists \mathbf{Q_p}, \mathbf{Q_c} \succeq \mathbf{0}, \ ((R_p, R_c), \mathbf{Q_p}, \mathbf{Q_c}) \in \mathcal{R}_{conv}^{\alpha, \boldsymbol{\Sigma_z}} \right\}. \tag{8}$$

Also, let $\mathcal{R}_{out}$ be represented as

$$\mathcal{R}_{out} = \bigcap_{\Sigma_z} \bigcap_{\alpha > 0} \mathcal{R}_{out}^{\alpha}. \tag{9}$$

Then, the next theorem describes an outer bound on the capacity region of the MCC.

Theorem 3.2: The capacity region of the MCC, $\mathcal{C}_{MCC}$ satisfies

$$\begin{array}{ll} \mathcal{C}_{MCC} & \subseteq \quad \mathcal{R}_{out}^{\alpha, \boldsymbol{\Sigma_z}}, \forall \alpha > 0, \boldsymbol{\Sigma_z} \\ \mathcal{C}_{MCC} & \subseteq \quad \mathcal{R}_{out}. \end{array} \tag{10}$$

The proof is given in Section V and proceeds by a series of channel transformations. Each channel transformation results in a new channel whose capacity region is in general a superset (outer bound) of the capacity region of the preceding channel.

Let $\text{BC}(\mathbf{H_1}, \mathbf{H_2}, P)$ denote a two user MIMO broadcast channel with channel matrices given by $\mathbf{H_1}$ and $\mathbf{H_2}$ and with a transmitter power constraint of P. Let $\mathcal{C}_{BC}^{H_1, H_2, P}$ denote the capacity region of $\text{BC}(\mathbf{H_1}, \mathbf{H_2}, P)$.

Let $\mathcal{R}_{part,conv}^{\alpha}$ denote the set described by

$$\left\{ \begin{array}{l} \left((R_p, R_c), \mathbf{Q_p}, \boldsymbol{\Sigma_{c,c}} \right) : \\ R_p \geq 0, R_c \geq 0, \mathbf{Q_p} \succeq \mathbf{0}, \boldsymbol{\Sigma_{c,c}} \succeq \mathbf{0}, \\ R_p \leq \log \left| \mathbf{I} + \mathbf{G}_\alpha \mathbf{Q_p} \mathbf{G}_\alpha^{\dagger} + \frac{1}{\alpha} \mathbf{H_{c,p}} \boldsymbol{\Sigma_{c,c}} \boldsymbol{\Sigma_{c,p}^{\dagger}} \right| \\ \qquad - \log \left| \mathbf{I} + \frac{1}{\alpha} \mathbf{H_{c,p}} \boldsymbol{\Sigma_{c,c}} \mathbf{H_{c,p}^{\dagger}} \right| \\ R_c \leq \log \left| \mathbf{I} + \frac{1}{\alpha} \mathbf{H_{c,c}} \boldsymbol{\Sigma_{c,c}} \mathbf{H_{c,c}^{\dagger}} \right| \\ \text{Tr}(\mathbf{Q_p}) + \text{Tr}(\boldsymbol{\Sigma_{c,c}}) \leq P_p + \alpha P_c \end{array} \right\}. \tag{11}$$

We let $\mathcal{R}_{part,out}^{\alpha}$ to denote the closure of the convex hull of the set of rate pairs described by

$$\left\{ \begin{array}{l} (R_p, R_c) : \exists \mathbf{Q_p}, \mathbf{\Sigma_{c,c}} \succeq \mathbf{0} \text{ and} \\ \quad ((R_p, R_c), \mathbf{Q_p}, \mathbf{\Sigma_{c,c}}) \in \mathcal{R}_{part,conv}^{\alpha} \end{array} \right\}. \quad (12)$$

Let $\mathbf{K} = [\mathbf{0} \ \mathbf{H_{c,c}}/\sqrt{\alpha}]$. We show that if the boundary of the rate region described by $\mathcal{R}_{part,out}^{\alpha}$ partially meets the boundary of the capacity region of $BC(\mathbf{G}_{\alpha}, \mathbf{K}, P_p + \alpha P_c)$, then the boundary of $\mathcal{R}_{part,out}^{\alpha}$ partially meets the boundary of the rate region described by $\mathcal{R}_{out}^{\alpha, \mathbf{\Sigma_z}}$ in (8) for some $\mathbf{\Sigma_z}$. We formally state the result in Theorem 3.3. For notational convenience, we will denote the capacity region of $BC(\mathbf{G}_{\alpha}, \mathbf{K}, P_p + \alpha P_c)$ by $\mathcal{C}_{BC}^{\alpha}$.

Theorem 3.3: Let $\mu \geq 1$ and $\alpha > 0$. If

$$\max_{(R_p, R_c) \in \mathcal{R}_{part,out}^{\alpha}} \mu R_p + R_c = \max_{(R_p, R_c) \in \mathcal{C}_{BC}^{\alpha}} \mu R_p + R_c, \quad (13)$$

then, we have

$$\max_{(R_p, R_c) \in \mathcal{R}_{part,out}^{\alpha}} \mu R_p + R_c = \inf_{\mathbf{\Sigma_z}} \max_{(R_p, R_c) \in \mathcal{R}_{out}^{\alpha, \mathbf{\Sigma_z}}} \mu R_p + R_c. \quad (14)$$

The proof of the theorem is described in Section V. Hence, if the condition (13) is satisfied, the rate region described by $\mathcal{R}_{part,out}^{\alpha}$ is an outer bound on the capacity region of the MCC in terms of maximizing the μ- sum $\mu R_p + R_c$.

Let $(\hat{R}_p, \hat{R}_c)$ be a point on the boundary of the capacity region $\mathcal{C}_{MCC}$. Then, there exists a $\mu \geq 0$ such that

$$(\hat{R}_p, \hat{R}_c) = \arg \max_{(R_p, R_c) \in \mathcal{C}_{MCC}} \mu R_p + R_c.$$

The next theorem shows that if (R_p, R_c) lies on the boundary of the achievable region given by $\mathcal{R}_{in}$, then (R_p, R_c) lies on the boundary of $\mathcal{R}_{part,out}^{\alpha}$ for some $\alpha > 0$. That is, the theorem describes conditions of optimality of the achievable region $\mathcal{R}_{in}$.

Theorem 3.4: For any $\mu > 0$,

$$\max_{(R_p, R_c) \in \mathcal{R}_{in}} \mu R_p + R_c = \inf_{\alpha > 0} \max_{(R_p, R_c) \in \mathcal{R}_{part,out}^{\alpha}} \mu R_p + R_c.$$

Also, there exists $\alpha^* \in (0, \infty)$, such that for any $\mu \geq 1$, $(R_{p,\mu}, R_{c,\mu}) = \arg \max_{(R_p, R_c) \in \mathcal{R}_{in}} \mu R_p + R_c$ is a point on the boundary of the capacity region of the MIMO cognitive channel if the condition given by (13) is satisfied for α^*.

The proof of the theorem is described in Section VI and is based on optimization techniques.

IV. ACHIEVABLE REGION

Proof of Theorem 3.1 : In this section, we show that the rate region $\mathcal{R}_{in}$ given by (4) is achievable on the MCC.

Encoding rule for Licensed user (E_p^n) : For every message $m_p \in \{1, \ldots, 2^{nR_p}\}$, the licensed encoder generates a n length codeword $\mathbf{X_p}^n(m_p)$, according to the distribution $p(\mathbf{X_p}^n) = $

$\Pi_{i=1}^n p(\mathbf{X_p}(i))$, and $X_p(i) \backsim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_p})$ such that $\mathbf{\Sigma_p} \succeq \mathbf{0}$ and $\text{Tr}(\mathbf{\Sigma_p}) \leq P_p$.

Encoding rule for the cognitive user (E_c^n): The cognitive encoder acts in two stages. For every message pair (m_p, m_c), the cognitive encoder first generates a codeword $\mathbf{X_{c,p}}^n(m_p, m_c)$ for the primary message m_p according to $\Pi_{i=1}^n p(\mathbf{X_{c,p}}(i) | \mathbf{X_p}(i))$, where $p(\mathbf{X_{c,p}}(i)) \backsim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_{c,p}})$ and the joint distribution of $(\mathbf{X_p}(i), \mathbf{X_{c,p}}(i))$ is given by

$$p(\mathbf{X_p}(i), \mathbf{X_{c,p}}(i)) \backsim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{\Sigma_p} & \mathbf{Q} \\ \mathbf{Q^{\dagger}} & \mathbf{\Sigma_{c,p}} \end{bmatrix}\right). \quad (15)$$

Here, $\mathbf{Q}$ denotes the correlation between $\mathbf{X_p}(i)$ and $\mathbf{X_{c,p}}(i)$. In the second stage, the cognitive encoder generates $\mathbf{X_{c,c}}^n$ which encodes message m_c. The codeword $\mathbf{X_{c,c}}^n$ is generated using Costa precoding [21] by treating $\mathbf{H_{p,p}X_p}^n + \mathbf{H_{c,c}X_{c,p}}^n$ as non causally known interference. A characteristic feature of Costa's precoding is that $\mathbf{X_{c,c}}^n$ is independent of $\mathbf{X_{c,p}}^n$, and $\mathbf{X_{c,c}}^n$ is distributed as $\Pi_{i=1}^n p(\mathbf{X_{c,c}}(i))$, where $\mathbf{X_{c,c}}(i) \backsim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_{c,c}})$. Note that the codeword $\mathbf{X_{c,p}}^n$ is used to convey message m_p to the licensed receiver and the codeword $\mathbf{X_{c,c}}^n$ is used to convey message m_c to the cognitive receiver. The two codewords $\mathbf{X_{c,p}}^n$ and $\mathbf{X_{c,c}}^n$ are superimposed to form the cognitive codeword $\mathbf{X_c}^n = \mathbf{X_{c,p}}^n + \mathbf{X_{c,c}}^n$. It is clear that $\mathbf{X_c}^n$ is distributed as $\Pi_{i=1}^n p(\mathbf{X_c}(i))$, $\mathbf{X_c}(i) \backsim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_c})$, where $\mathbf{\Sigma_c} = \mathbf{\Sigma_{c,p}} + \mathbf{\Sigma_{c,c}}$. The covariance matrices satisfy the constraints $\mathbf{\Sigma_{c,p}} \succeq \mathbf{0}, \mathbf{\Sigma_{c,c}} \succeq \mathbf{0}, \text{Tr}(\mathbf{\Sigma_c}) \leq P_c$.

Decoding rule for the licensed receiver (D_p^n) : The licensed receiver receives $\mathbf{H_{p,p}X_p}^n + \mathbf{H_{c,p}}(\mathbf{X_{c,p}}^n + \mathbf{X_{c,c}}^n) + \mathbf{Z_p}^n$. It treats $\mathbf{H_{p,p}X_p}^n + \mathbf{H_{c,p}X_{c,p}}^n$ as the valid codeword and $\mathbf{H_{c,p}X_{c,c}}^n + \mathbf{Z_p}^n$ as Gaussian noise. Taking $\mathbf{G} = [\mathbf{H_{p,p}} \ \mathbf{H_{c,p}}]$ and $\mathbf{X_{p,net}}^n = \begin{bmatrix} \mathbf{X_p}^n \\ \mathbf{X_{c,p}}^n \end{bmatrix}$, the received vector at the licensed receiver is

$$\mathbf{Y_p}^n = \mathbf{G X_{p,net}}^n + \mathbf{H_{c,p} X_{c,c}}^n + \mathbf{Z_p}^n. \quad (16)$$

The covariance matrix of $\mathbf{X_{p,net}}$ is denoted by $\mathbf{\Sigma_{p,net}} = \begin{bmatrix} \mathbf{\Sigma_p} & \mathbf{Q} \\ \mathbf{Q^{\dagger}} & \mathbf{\Sigma_{c,p}} \end{bmatrix}$, where $\mathbf{Q} = E[\mathbf{X_p X_{c,p}^{\dagger}}]$. In this setup, we use steps identical to that used for MIMO channel with colored noise in [20, Section 9.5] to show that, for any $\epsilon > 0$, there exists a block length n_1 so that for any $n \geq n_1$, the licensed decoder can recover the message m_p with probability of error $< \epsilon$ if

$$R_p \leq \log \left| \mathbf{I} + \mathbf{G \Sigma_{p,net} G^{\dagger}} + \mathbf{H_{c,p} \Sigma_{c,c} H_{c,p}^{\dagger}} \right| \\ - \log \left| \mathbf{I} + \mathbf{H_{c,p} \Sigma_{c,c} H_{c,p}^{\dagger}} \right|. \quad (17)$$

Decoding rule for the cognitive user (D_c^n) : The cognitive decoder is the Costa decoder (with the knowledge of the encoder, E_c^n). The cognitive receiver receives $\mathbf{Y_c}^n = \mathbf{H_{p,c} X_p}^n + \mathbf{H_{c,c}}(\mathbf{X_{c,p}}^n + \mathbf{X_{c,c}}^n) + \mathbf{Z_c}^n$. Here, the non-causally known interference $\mathbf{H_{p,c} X_p}^n + \mathbf{H_{c,c} X_{c,p}}^n$ is canceled by the Costa precoder. To show this formally, we follow steps similar to Eqns (3) to (7) in [21]. We get that, for any $\epsilon_2 > 0$, there exists n_2 such that for $n \geq n_2$, the cognitive decoder can recover the message m_c with probability of error $< \epsilon_2$ if

$$R_c \leq \log \left| \mathbf{I} + \mathbf{H_{c,c} \Sigma_{c,c} H_{c,c}^{\dagger}} \right|. \quad (18)$$

Note that the achievable scheme holds for all possible covariance matrices $\mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{\Sigma_{c,c}}$ that are positive semidefinite and satisfy the power constraints $\mathrm{Tr}(\mathbf{\Sigma_p}) \leq P_p, \mathrm{Tr}(\mathbf{\Sigma_{c,p}} + \mathbf{\Sigma_{c,c}}) \leq P_c$. Hence, $\mathcal{R}_{in}$, which is the set of all achievable rate pairs described by (4), is achievable for any code length $n \geq \max(n_1, n_2)$. ∎

V. OUTER BOUND ON THE CAPACITY REGION

In this section, we prove that the rate region described by $\mathcal{R}_{out}^{\alpha, \mathbf{\Sigma_z}}$ is an outer bound on the capacity region of the Gaussian MIMO cognitive channel. The proof proceeds by a series of channel transformations where each transformation creates an outer bound on the channel at the previous stage. At the final stage, we obtain a physically degraded broadcast channel. The capacity region of this channel is now known [17][18][19] and is used as the outer bound for the capacity region of the MIMO cognitive channel. Figure 2 depicts the various channel configurations considered, and the system equations of all the configurations. $\hat{\mathbf{Z}}_{\mathbf{p}}^n$ shown in Figures 2c, 2d and 2e has the same distribution as $\mathbf{Z_p}^n$, but has an arbitrary correlation with $\mathbf{Z_c}^n$. Before proving Theorem 3.2, we prove the following lemmas.

Transformation 1 (MIMO Cognitive Channel (MCC) → Scaled MIMO cognitive channel) : The scaled MIMO cognitive channel is defined in Figure 2b and Figure 3. In this transformation, the channel matrices $\mathbf{H_{c,p}}$ and $\mathbf{H_{c,c}}$ are scaled by $1/\sqrt{\alpha}$. Also, the power constraint at the cognitive transmitter is changed to αP_c.

Lemma 5.1: The capacity region of the MIMO cognitive channel is equal to the capacity region of the scaled MIMO cognitive channel (SMCC) for any $0 < \alpha < \infty$.

Proof : Let (R_p, R_c) be a rate pair that is achievable on the MCC. That is, for all $\epsilon_1, \epsilon_2 > 0$, there exists a n and a sequence of encoder decoder pairs at the licensed and cognitive transmitter and receiver $(E_p^n : m_p \to \mathbf{X_p}^n, D_p^n : \mathbf{Y_p}^n \to \hat{m}_p, E_c^n : (m_p, m_c) \to \mathbf{X_c}^n, D_c^n : \mathbf{Y_c}^n \to \hat{m}_c)$ such that the codewords $\mathbf{X_p}^n$ and $\mathbf{X_c}^n$ satisfy the power constraints given by (2) and the probability of decoding error is small $(Pr(m_p \neq \hat{m}_p) \leq \epsilon_1, Pr(m_c \neq \hat{m}_c) \leq \epsilon_2)$. We use the following encoder decoder pairs at the licensed and cognitive transmitters and receivers of the scaled MIMO cognitive channel. $E_p^n : m_p \to \mathbf{X_p}^n, D_p^n : \mathbf{Y_p}^n \to \hat{m}_p, E_c^n : (m_p, m_c) \to \sqrt{\alpha}\mathbf{X_c}^n, D_c^n : \mathbf{Y_c}^n \to \hat{m}_c$. It follows that using these encoder and decoder pairs, the licensed and cognitive codewords satisfy the new power constraints of P_p and αP_c respectively. Also, the system equation is the same as that of the MCC and $Pr(m_p \neq \hat{m}_p) \leq \epsilon_1$ and $Pr(m_c \neq \hat{m}_c) \leq \epsilon_2$. Hence, the rate pair (R_p, R_c) is achievable on the scaled MIMO cognitive channel. Hence, the capacity region of the SMCC is a superset of the capacity region of the MCC.

Similarly, we can also establish this in the other direction, namely we can treat the MCC as the scaled version of the SMCC (scaling by $1/\alpha$). Therefore, it can be shown that the capacity region of the MCC is a superset of the capacity region of the SMCC.

Hence, the capacity region of the MCC is equal to the capacity region of the SMCC. ∎

Transformation 2 (scaled MIMO cognitive channel (SMCC) → scaled MIMO cognitive channel A (SMCCA)) : The scaled MIMO cognitive channel A (SMCCA) is described in Figure 2c and Figure 4. In this transformation, we provide a modified version of $\mathbf{Y_p}^n$, which is $\hat{\mathbf{Y}}_{\mathbf{p}}^n$ to the cognitive receiver. $\hat{\mathbf{Y}}_{\mathbf{p}}^n$ is corrupted by noise $\hat{\mathbf{Z}}_{\mathbf{p}}^n$, which has the same probability distribution as that of $\mathbf{Z_p}^n$ (i.e., complex Gaussian with zero mean and identity covariance matrix), but is permitted to be correlated with $\mathbf{Z_p}^n$ or $\mathbf{Z_c}^n$. In fact, we assume that the joint probability distribution of $(\hat{\mathbf{Z}}_{\mathbf{p}}(i), \mathbf{Z_c}(i))$ is given by

$$p(\hat{\mathbf{Z}}_{\mathbf{p}}(i), \mathbf{Z_c}(i)) = \mathcal{N}(0, \mathbf{\Sigma_z}), \qquad (19)$$

where $\mathbf{\Sigma_z}$ has the form given by (6). The received vector $\hat{\mathbf{Y}}_{\mathbf{p}}^n$ is made available to the cognitive receiver by transforming the channel matrices $\mathbf{H_{p,c}}$ and $\mathbf{H_{c,c}}/\sqrt{\alpha}$ to $K_1 = \begin{bmatrix} \mathbf{H_{p,p}} \\ \mathbf{H_{p,c}} \end{bmatrix}$ and $K_2 = \begin{bmatrix} \mathbf{H_{c,p}}/\sqrt{\alpha} \\ \mathbf{H_{c,c}}/\sqrt{\alpha} \end{bmatrix}$ respectively. Hence, the received vector at the cognitive receiver is $\begin{bmatrix} \hat{\mathbf{Y}}_{\mathbf{p}}^n \\ \mathbf{Y_c}^n \end{bmatrix}$.

Lemma 5.2: The capacity region of the scaled MIMO cognitive channel A (SMCCA) is a superset of the capacity region of the scaled MIMO cognitive channel (SMCC).

Proof : Let the rate pair (R_p, R_c) be achievable on the SMCC. That is, for all $\epsilon_1, \epsilon_2 > 0$, there exists a n and a sequence of encoder decoder pairs at the licensed and cognitive transmitter and receiver $(E_p^n : m_p \to \mathbf{X_p}^n, D_p^n : \mathbf{Y_p}^n \to \hat{m}_p, E_c^n : (m_p, m_c) \to \mathbf{X_c}^n, D_c^n : \mathbf{Y_c}^n \to \hat{m}_c)$ such that the codewords $\mathbf{X_p}^n$ and $\mathbf{X_c}^n$ satisfy the power constraints and the probability of decoding error is small $(Pr(m_p \neq \hat{m}_p) \leq \epsilon_1, Pr(m_c \neq \hat{m}_c) \leq \epsilon_2)$. In the SMCCA, we can use the same encoder decoder pair E_p^n and D_p^n at the licensed transmitter and receiver to achieve a rate R_p with probability of decoding error $< \epsilon_1$. Also, by ignoring the received vector $\hat{\mathbf{Y}}_{\mathbf{p}}^n$ at the cognitive receiver, we can use E_c^n and D_c^n at the cognitive transmitters and receivers to achieve a rate R_c with the decoding probability of error $< \epsilon_2$. Hence, the rate pair (R_p, R_c) is achievable on the scaled MIMO cognitive channel A (SMCCA). Therefore, the capacity region of the SMCCA is a superset of the capacity region of the SMCC. ∎

Transformation 3 (scaled MIMO cognitive channel A (SMCCA) → scaled MIMO cognitive channel B (SMCCB)) : The scaled MIMO cognitive channel (B) is described in Figure 2d and Figure 5. The channel matrix from the licensed transmitter to the cognitive receiver is modified from $\mathbf{K_1} = \begin{bmatrix} \mathbf{H_{p,p}} \\ \mathbf{H_{p,c}} \end{bmatrix}$ to $\overline{\mathbf{K_1}} = \begin{bmatrix} \mathbf{H_{p,p}} \\ \mathbf{0} \end{bmatrix}$. Hence, the received vector at the cognitive receiver is given by $\begin{bmatrix} \hat{\mathbf{Y}}_{\mathbf{p}}^n \\ \overline{\mathbf{Y_c}^n} \end{bmatrix}$ where $\overline{\mathbf{Y_c}^n} = \frac{\mathbf{H_{c,c}}}{\sqrt{\alpha}}\mathbf{X_c}^n + \mathbf{Z_c}^n$. The intuition behind the transformation is to remove the
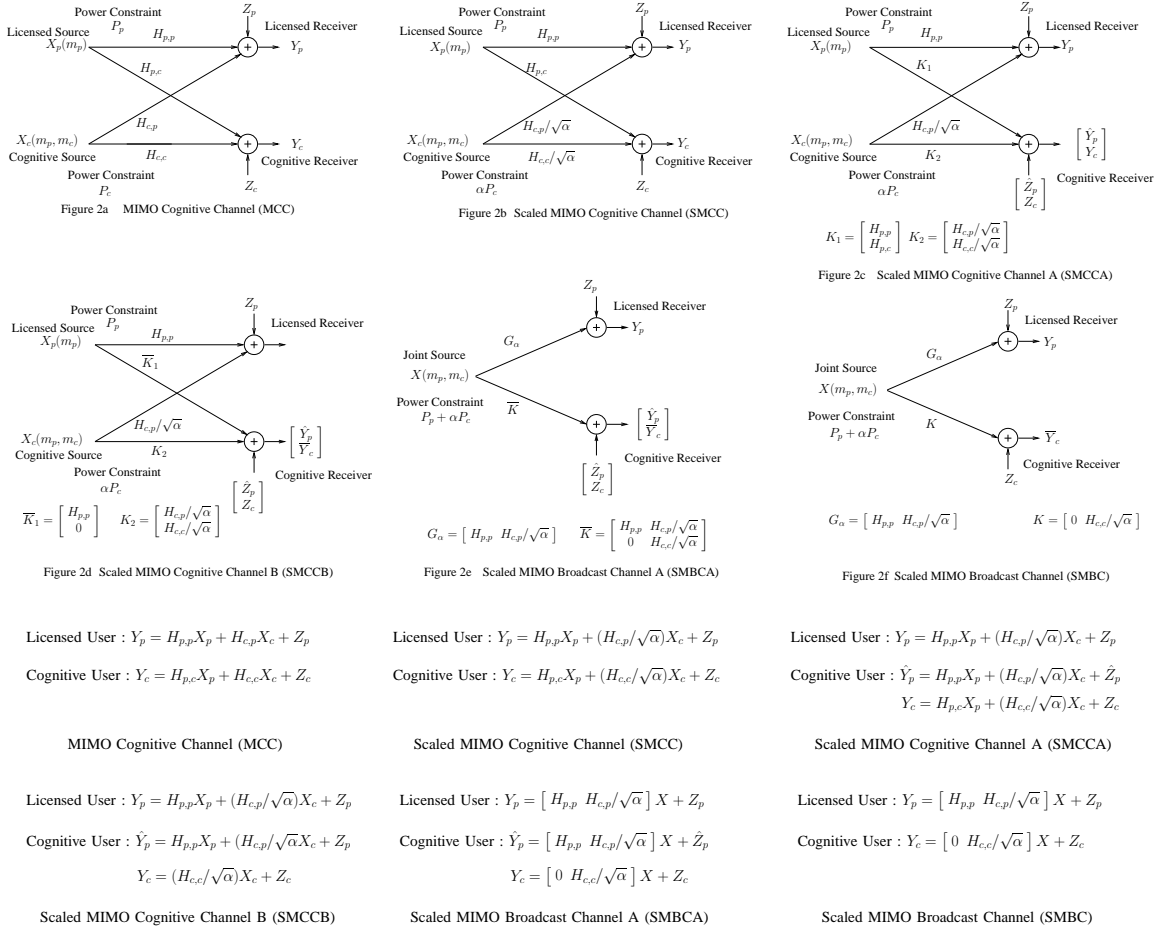
Fig. 2. Channel Configurations and their System Equations
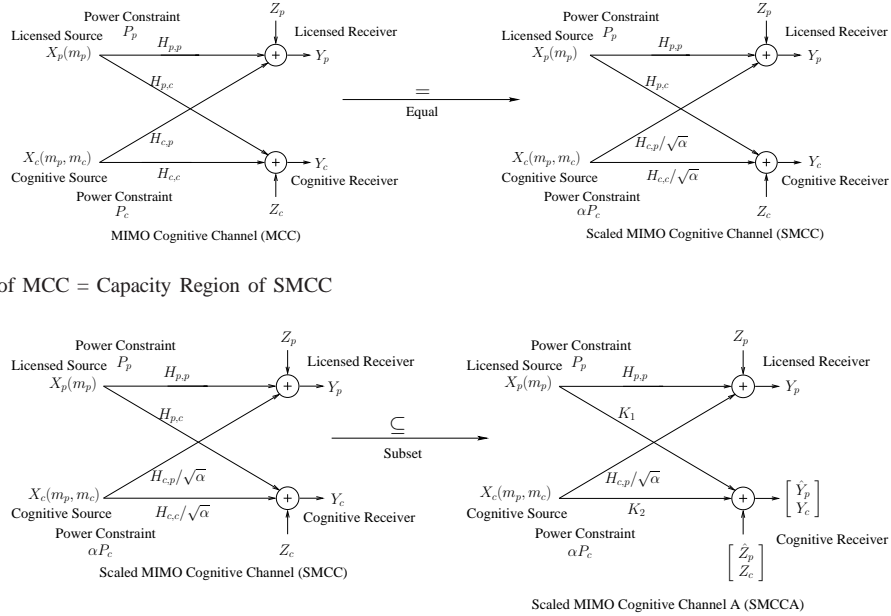
Fig. 3. Capacity Region of MCC = Capacity Region of SMCC
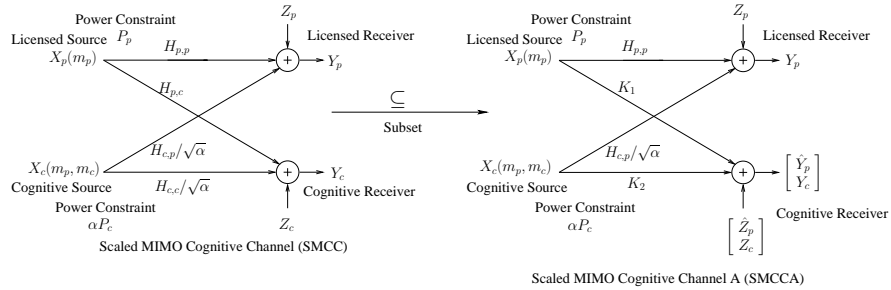
Fig. 4. Capacity Region of SMCC $\subseteq$ Capacity Region of SMCCA

original interference caused by the licensed transmitter to the cognitive receiver.

Lemma 5.3: The capacity region of the scaled MIMO cogni-

tive channel B (SMCCB) is equal to the capacity region of the scaled MIMO cognitive channel A (SMCCA).

Proof : Let the rate pair (R_p, R_c) be achievable on the

Fig. 5.   Capacity Region of SMCCA = Capacity Region of SMCCB

SMCCA. This implies that for every $\epsilon_1, \epsilon_2 > 0$, there exists encoder-decoder pair for the licensed user $(E_p^n(\epsilon_1), D_p^n(\epsilon_1))$ and for the cognitive user $(E_c^n(\epsilon_2), D_c^n(\epsilon_2))$ such that the probability of decoding error is less than ϵ_1 and ϵ_2 respectively for the licensed and cognitive user. Let $\delta_1, \delta_2 \in (0,1)$. In SMCCB, the licensed 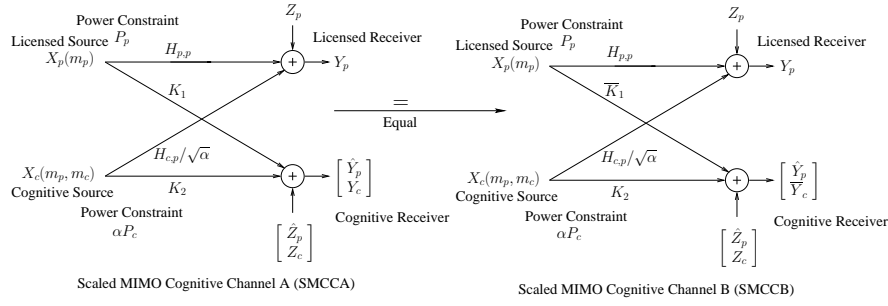user can employ $E_p^n(\min(\delta_1/2, \delta_2/2))$, $D_p^n(\min(\delta_1/2, \delta_2/2))$ to decode m_p with a probability of error $\leq \delta_1/2 < \delta_1$. The cognitive receiver uses $E_p^n(\min(\delta_1/2, \delta_2/2))$, $D_p^n(\min(\delta_1/2, \delta_2/2))$ on $\hat{\mathbf{Y}}_{\mathbf{p}}^n$ to obtain m_p with probability of error $\leq \delta_1/2$. The cognitive receiver can now construct $\mathbf{X_p}^n$ and hence $\mathbf{H_{p,c}X_p}^n$. Thus, the cognitive receiver recovers $\mathbf{Y_c}^n = \mathbf{H_{p,c}X_p}^n + \frac{\mathbf{H_{c,c}}}{\sqrt{\alpha}}\mathbf{X_{c,c}^n} + \mathbf{Z_c}^n$. Now, it uses, $E_c^n(\delta_2/2), D_c^n(\delta_2/2)$ to obtain m_c with probability of error $\leq \delta_2/2$. Clearly, the probability of error in recovering m_c is less than δ_2. Hence, the rate pair (R_p, R_c) is achievable on SMCCB. Therefore, the capacity region of SMCCB is a superset of the capacity region of SMCCA.

Let the rate pair (R_p, R_c) be achievable on SMCCB. Then, for every $\epsilon_1, \epsilon_2 > 0$, there exists encoder-decoder pair for the licensed user $(E_p^n(\epsilon_1), D_p^n(\epsilon_1))$ and for the cognitive user $(E_c^n(\epsilon_2), D_c^n(\epsilon_2))$ such that the probability of decoding error is less than ϵ_1 and ϵ_2 respectively for the licensed and cognitive user. Let $\delta_1, \delta_2 > 0$. In SMCCA, the licensed user can employ $E_p^n(\min(\delta_1/2, \delta_2/2))$, $D_p^n(\min(\delta_1/2, \delta_2/2))$ to decode m_p with a probability of error $\leq \delta_1/2 < \delta_1$. The cognitive user employs $E_p^n(\min(\delta_1/2, \delta_2/2))$, $D_p^n(\min(\delta_1/2, \delta_2/2))$ on $\hat{\mathbf{Y}}_{\mathbf{p}}^n$ to obtain m_p with probability of error $\leq \delta_2/2$. The cognitive receiver can now construct $\mathbf{X_p}^n$ and hence $\mathbf{H_{p,c}X_p}^n$. Hence, the cognitive receiver subtracts $\mathbf{H_{p,c}X_p}^n$ from $\mathbf{Y_c}^n$ to obtain $\overline{\mathbf{Y}}_{\mathbf{c}}^{\,n}$. The cognitive receiver can now use $E_c^n(\delta_2/2), D_c^n(\delta_2/2)$ to obtain m_c with probability of error $< \delta_2$. Thus, the rate pair (R_p, R_c) is achievable on SMCCA.

Therefore, the capacity region of the SMCCA is equal to the capacity region of the SMCCB. ∎

Transformation 4 (scaled MIMO cognitive channel (B) → scaled MIMO broadcast channel A (SMBCA)): The scaled MIMO broadcast channel A (SMBCA) is depicted in Figure 2e and Figure 6. We let the two transmitters to cooperate and transform it into a broadcast channel with a sum power constraint of $P_p + \alpha P_c$. The new channel matrices from the combined transmitters to the licensed and cognitive receivers are given by $\mathbf{G}_\alpha = \begin{bmatrix} \mathbf{H_{p,p}} & \mathbf{H_{c,p}}/\sqrt{\alpha} \end{bmatrix}$ and

$\overline{\mathbf{K}} = \begin{bmatrix} \mathbf{H_{p,p}} & \mathbf{H_{c,p}}/\sqrt{\alpha} \\ \mathbf{0} & \mathbf{H_{c,c}}/\sqrt{\alpha} \end{bmatrix}$ respectively.

Lemma 5.4: The capacity region of the scaled MIMO broadcast channel A (SMBCA) is a superset of the capacity region of scaled MIMO cognitive channel B (SMCCB).

Proof : Let the rate pair (R_p, R_c) be achievable on the SMCCB. In the SMBCA, using no collaboration between the two transmitters and using separate power constraints of P_p and αP_c respectively, we reduce the SMBCA to the SMCCB. Hence, the rate pair (R_p, R_c) is achievable on the SMBCA. Therefore, the capacity region of the SMBCA is a superset of the capacity region of the SMCCB. ∎

We have showed that for any $\alpha > 0$, $\mathcal{C}_{MCC} = \mathcal{C}_{SMCC} \subseteq \mathcal{C}_{SMCCA} = \mathcal{C}_{SMCCB} \subseteq \mathcal{C}_{SMBCA}$. Hence, the capacity region of the scaled MIMO broadcast channel A (SMBCA) is a superset of the capacity region of the MIMO cognitive channel (MCC).

Proof of Theorem 3.2 : In the SMBCA, let $\mathbf{Q_p}$ denote the covariance matrix of the codeword for the licensed user and let $\mathbf{Q_c}$ denote the covariance matrix for the cognitive user. The SMBCA is a physically degraded broadcast channel. Hence, the capacity region of the SMBCA (as given by [17]) denoted by $\mathcal{C}_{SMBCA}$ is the closure of the convex hull of the set of rate pairs described by

$$\left\{ \begin{array}{l} (R_p, R_c) : R_p \geq 0, R_c \geq 0 \\ R_p \leq \log \left| \mathbf{I} + \mathbf{G}_\alpha \mathbf{Q_p} \mathbf{G}_\alpha^\dagger + \mathbf{G}_\alpha \mathbf{Q_c} \mathbf{G}_\alpha^\dagger \right| \\ \qquad - \log \left| \mathbf{I} + \mathbf{G}_\alpha \mathbf{Q_c} \mathbf{G}_\alpha^\dagger \right| \\ R_c \leq \log \left| \mathbf{\Sigma_z} + \overline{\mathbf{K}} \mathbf{Q_c} \overline{\mathbf{K}^\dagger} \right| - \log |\mathbf{\Sigma_z}| \\ \forall \mathbf{Q_p} \succeq \mathbf{0}, \mathbf{Q_c} \succeq \mathbf{0} \\ \mathrm{Tr}(\mathbf{Q_p}) + \mathrm{Tr}(\mathbf{Q_c}) \leq P_p + \alpha P_c \end{array} \right\}. \quad (20)$$

Also, this is the outer bound of the MCC. Hence, $\mathcal{R}_{out}^{\alpha, \mathbf{\Sigma_z}}$ described by (8) is an outer bound on the capacity region of the MCC. Hence, $\mathcal{C}_{MCC} \subseteq \mathcal{R}_{out}^{\alpha, \mathbf{\Sigma_z}}$. Also, $\mathcal{C}_{MCC} \subseteq \mathcal{R}_{out}$, where $\mathcal{R}_{out}$ is described in (9). ∎

Transformation 5 (scaled MIMO broadcast channel A (SMBCA) → scaled MIMO broadcast channel (SMBC)) : The scaled MIMO broadcast channel (SMBC) is depicted in Figure 2f and Figure 7. We change the received vector at the cognitive
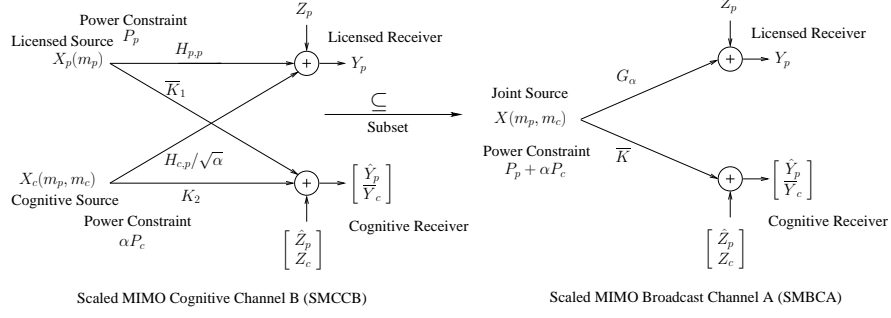
Fig. 6. Capacity Region of SMCCB $\subseteq$ Capacity Region of SMBCA

receiver from $\left[\begin{array}{c} \hat{\mathbf{Y}}_{\mathbf{p}}^{n} \\ \overline{\mathbf{Y}_{\mathbf{c}}}^{n} \end{array}\right]$ to $\overline{\mathbf{Y}_{\mathbf{c}}}^{n}$. This is done by changing the channel matrix from the joint transmitters to the cognitive receiver to $\mathbf{K} = \left[\begin{array}{cc} \mathbf{0} & \mathbf{H}_{\mathbf{c,c}}/\sqrt{\alpha} \end{array}\right]$.

Lemma 5.5 ([23]): The capacity region of the SMBCA is a superset of the capacity region of the scaled MIMO broadcast channel (SMBC).

Proof: Let the rate pair (R_p, R_c) be achievable on the SMBC. That is, for all $\epsilon_1, \epsilon_2 > 0$, there exists a n and a sequence of encoder decoder pairs at the transmitter and the two receivers $(E^n : (m_p, m_c) \to \mathbf{X}^n, D_p^n : \mathbf{Y_p}^n \to \hat{m}_p, D_c^n : \mathbf{Y_c}^n \to \hat{m}_c)$ such that the codeword $\mathbf{X}^n$ satisfies the power constraint of $P_p + \alpha P_c$ and the probability of decoding error is small $(Pr(m_p \neq \hat{m}_p) \leq \epsilon_1, Pr(m_c \neq \hat{m}_c) \leq \epsilon_2)$.

In the SMBCA, the transmitter and the receivers use the same coding strategy. The licensed receiver can decode message m_p at a rate R_p. The cognitive receiver can ignore $\hat{\mathbf{Y}}_{\mathbf{p}}^n$ and use just $\overline{\mathbf{Y}_{\mathbf{c}}}^n$ to decode message m_c at a rate R_c. Hence, the rate pair (R_p, R_c) is achievable in the SMBCA. Hence, the capacity region of the SMBCA is in general a superset of the capacity region of the SMBC. ∎

We describe one more lemma whose result will be used in the proof of Theorem (3.3).

Lemma 5.6 ([23]): Let $\mathcal{C}_{SMBC}$ denote the capacity region of the scaled MIMO broadcast channel described in Figure 2f. Then, for any $\mu \geq 1$,

$$\sup_{(R_p, R_c) \in \mathcal{C}_{SMBC}} \mu R_p + R_c = \inf_{\boldsymbol{\Sigma}_{\mathbf{z}}} \sup_{(R_p, R_c) \in \mathcal{C}_{SMBCA}} \mu R_p + R_c.$$

The proof is described in [23, Section 5.1] and is omitted here.

We now give the proof for Theorem (3.3).

Proof of Theorem 3.3 : It was shown in [17] that Gaussian codebooks (i.e., codebooks generated using i.i.d. realizations of an appropriate Gaussian random variable) achieve the capacity region for the MIMO broadcast channel. In SMBC, let $\mathbf{Q_p}$ denote the covariance of codeword $\mathbf{X}^n$ for the licensed user and $\mathbf{Q_c}$ denote the covariance matrix for the cognitive user. The covariance matrices satisfy the joint power constraint

$\text{Tr}(\mathbf{Q_p} + \mathbf{Q_c}) \leq P_p + \alpha P_c$. Let $\mathcal{R}_{SMBC,1}^{\alpha}$ denote the closure of the convex hull of the set of rate pairs described by

$$\left\{\begin{array}{l} (R_p, R_c) : R_p \geq 0, R_c \geq 0 \\ R_p \leq \log\left|\mathbf{I} + \mathbf{G}_\alpha \mathbf{Q_p} \mathbf{G}_\alpha^\dagger + \mathbf{G}_\alpha \mathbf{Q_c} \mathbf{G}_\alpha^\dagger\right| \\ \quad - \log\left|\mathbf{I} + \mathbf{G}_\alpha \mathbf{Q_c} \mathbf{G}_\alpha^\dagger\right| \\ R_c \leq \log\left|\mathbf{I} + \mathbf{K} \mathbf{Q_c} \mathbf{K}^\dagger\right| \\ \forall \mathbf{Q_p} \succeq \mathbf{0}, \mathbf{Q_c} \succeq \mathbf{0} \\ \text{Tr}(\mathbf{Q_p}) + \text{Tr}(\mathbf{Q_c}) \leq P_p + \alpha P_c \end{array}\right\}. \quad (21)$$

Similarly, let $\mathcal{R}_{SMBC,2}^{\alpha}$ denote the closure of the convex hull of the set of rate pairs described by

$$\left\{\begin{array}{l} (R_p, R_c) : R_p \geq 0, R_c \geq 0 \\ R_p \leq \log\left|\mathbf{I} + \mathbf{G}_\alpha \mathbf{Q_p} \mathbf{G}_\alpha^\dagger\right| \\ R_c \leq \log\left|\mathbf{I} + \mathbf{K} \mathbf{Q_p} \mathbf{K}^\dagger + \mathbf{K} \mathbf{Q_c} \mathbf{K}^\dagger\right| \\ \quad - \log\left|\mathbf{I} + \mathbf{K} \mathbf{Q_p} \mathbf{K}^\dagger\right| \\ \forall \mathbf{Q_p} \succeq \mathbf{0}, \mathbf{Q_c} \succeq \mathbf{0}, \\ \text{Tr}(\mathbf{Q_p}) + \text{Tr}(\mathbf{Q_c}) \leq P_p + \alpha P_c \end{array}\right\}. \quad (22)$$

The capacity region of SMBC, $\mathcal{C}_{SMBC}$ is the closure of the convex hull of $\mathcal{R}_{SMBC,1}^{\alpha} \cup \mathcal{R}_{SMBC,2}^{\alpha}$. That is,

$$\mathcal{C}_{SMBC} = \text{Cl}(\text{Co}(\mathcal{R}_{SMBC,1}^{\alpha} \cup \mathcal{R}_{SMBC,2}^{\alpha})). \quad (23)$$

$\mathcal{R}_{SMBC,1}^{\alpha}$ denotes the portion of the capacity region of SMBC where the licensed user's message is encoded first. That is, the cognitive receiver sees no interference. Hence, for $\mu \geq 1$, we have

$$\max_{(R_p, R_c) \in \mathcal{R}_{SMBC,1}^{\alpha}} \mu R_p + R_c = \max_{(R_p, R_c) \in \mathcal{C}_{SMBC}} \mu R_p + R_c.$$

Therefore, from Lemma 5.6, we have that for $\mu \geq 1$,

$$\max_{(R_p, R_c) \in \mathcal{R}_{SMBC,1}^{\alpha}} \mu R_p + R_c = \inf_{\boldsymbol{\Sigma}_{\mathbf{z}}} \max_{(R_p, R_c) \in \mathcal{C}_{SMBCA}} \mu R_p + R_c.$$

We can see that, $\mathcal{R}_{part,out}^{\alpha}$ described in (12) is a subset of $\mathcal{R}_{SMBC,1}^{\alpha}$ formed by restricting the covariance matrix $\mathbf{Q_c}$ to have the form

$$\mathbf{Q_c} = \left[\begin{array}{cc} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\mathbf{c,c}} \end{array}\right].$$

It can also be seen that $\mathcal{R}_{out}^{\alpha, \boldsymbol{\Sigma}_{\mathbf{z}}}$ described in (8) equals $\mathcal{C}_{SMBCA}$. Hence, it follows that for any $\mu \geq 1$ and for $\alpha > 0$, if

$$\max_{(R_p, R_c) \in \mathcal{R}_{part,out}^{\alpha}} \mu R_p + R_c = \max_{(R_p, R_c) \in \mathcal{C}_{BC}^{\alpha}} \mu R_p + R_c,$$
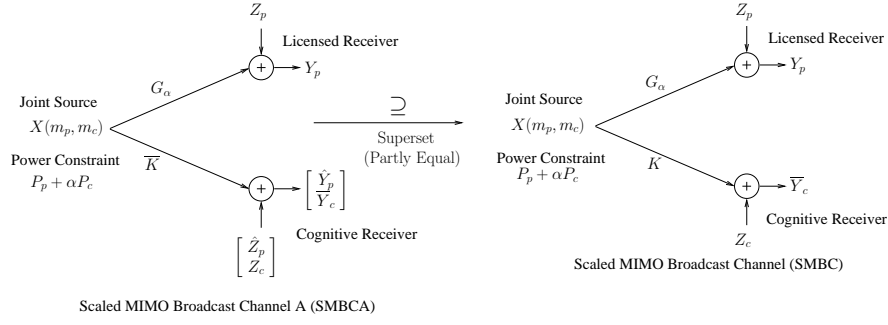
Fig. 7. Capacity Region of SMBCA $\supseteq$ Capacity Region of SMBC

then we have that

$$\max_{(R_p,R_c)\in\mathcal{R}^\alpha_{part,out}} \mu R_p + R_c = \inf_{\boldsymbol{\Sigma_z}} \max_{(R_p,R_c)\in\mathcal{R}^{\alpha,\boldsymbol{\Sigma_z}}_{out}} \mu R_p + R_c.$$

■

VI. OPTIMALITY OF THE ACHIEVABLE REGION

In this section, we describe conditions under which the achievable region described by $\mathcal{R}_{in}$ in (4) is optimal for a portion of the capacity region. In particular, we show that if (R_p, R_c) lies on the boundary of the achievable region given by $\mathcal{R}_{in}$, then (R_p, R_c) lies on the boundary of $\mathcal{R}^\alpha_{part,out}$ given by (12) for some $\alpha > 0$. That is, for any $\mu > 0$,

$$\sup_{(R_p,R_c)\in\mathcal{R}_{in}} \mu R_p + R_c = \inf_{\alpha>0} \sup_{(R_p,R_c)\in\mathcal{R}^\alpha_{part,out}} \mu R_p + R_c.$$

Then there exists $\alpha^* \in (0,\infty)$ such that, for any $\mu \geq 1$, $(R_{p,\mu}, R_{c,\mu}) = \arg\max_{(R_p,R_c)\in\mathcal{R}_{in}} \mu R_p + R_c$ is a point on the boundary of the capacity region of the MIMO cognitive channel if the condition (13) is satisfied for α^*.

We denote by $\mathcal{R}_{ach,rate}$, the set of all $((R_p, R_c), \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \mathbf{Q})$ given by

$$\left\{ \begin{array}{l} \left((R_p, R_c), \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \mathbf{Q} \right) : \\ R_p \geq 0, R_c \geq 0, \boldsymbol{\Sigma_p} \succeq \mathbf{0}, \boldsymbol{\Sigma_{c,p}} \succeq \mathbf{0}, \boldsymbol{\Sigma_{c,c}} \succeq \mathbf{0} \\ R_p \leq \log\left|\mathbf{I} + \mathbf{G}\boldsymbol{\Sigma_{p,net}}\mathbf{G}^\dagger + \mathbf{H_{c,p}}\boldsymbol{\Sigma_{c,c}}\mathbf{H}^\dagger_{\mathbf{c,p}}\right| \\ \qquad - \log\left|\mathbf{I} + \mathbf{H_{c,p}}\boldsymbol{\Sigma_{c,c}}\mathbf{H}^\dagger_{\mathbf{c,p}}\right| \\ R_c \leq \log\left|\mathbf{I} + \mathbf{H_{c,c}}\boldsymbol{\Sigma_{c,c}}\mathbf{H}^\dagger_{\mathbf{c,c}}\right| \\ \boldsymbol{\Sigma_{p,net}} = \left(\begin{array}{cc} \boldsymbol{\Sigma_p} & \mathbf{Q} \\ \mathbf{Q}^\dagger & \boldsymbol{\Sigma_{c,p}} \end{array} \right) \succeq \mathbf{0} \end{array} \right\}. \quad (24)$$

The rate pair that maximizes $\mu R_p + R_c$ in the achievable region is given by solving the optimization problem

$$\sup_{((R_p,R_c),\boldsymbol{\Sigma_p},\boldsymbol{\Sigma_{c,p}},\boldsymbol{\Sigma_{c,c}},\mathbf{Q})} \mu R_p + R_c \quad (25)$$

$$\text{such that} \quad ((R_p, R_c), \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \mathbf{Q})$$
$$\in \mathcal{R}_{ach,rate}$$
$$\text{Tr}(\boldsymbol{\Sigma_p}) \leq P_p$$
$$\text{Tr}(\boldsymbol{\Sigma_{c,p}} + \boldsymbol{\Sigma_{c,c}}) \leq P_c.$$

We define the functions $L(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \lambda_1, \lambda_2)$ and $g(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}})$ as follows

$$L(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \lambda_1, \lambda_2) = $$
$$\mu R_p + R_c - \lambda_1(\text{Tr}(\boldsymbol{\Sigma_p}) - P_p) - \quad (26)$$
$$\lambda_2(\text{Tr}(\boldsymbol{\Sigma_{c,p}} + \boldsymbol{\Sigma_{c,c}}) - P_c))$$

$$g(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}) = $$
$$\min_{\lambda_1\geq0,\lambda_2\geq0} L(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \lambda_1, \lambda_2). \quad (27)$$

The optimization problem given by

$$\max_{(R_p,R_c,\boldsymbol{\Sigma_p},\boldsymbol{\Sigma_{c,p}},\boldsymbol{\Sigma_{c,c}},\mathbf{Q})} g(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}) \quad (28)$$

$$\text{such that} \quad ((R_p, R_c), \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \mathbf{Q})$$
$$\in \mathcal{R}_{ach,rate}$$

has the same optimum value as that of (25). This is formally stated in the lemma below.

Lemma 6.1: Let M denote the optimal value of the optimization problem defined in (25), and U denote the optimal value of the optimization problem defined in (28). Then, $M = U$.

Proof : We show that for any set of covariance matrices $(\boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}})$ that do not satisfy the power constraints given by (2), $g(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}) = -\infty$. The power constraints can be violated by three means :

- $\text{Tr}(\boldsymbol{\Sigma_p}) > P_p$ and $\text{Tr}(\boldsymbol{\Sigma_{c,p}}) + \text{Tr}(\boldsymbol{\Sigma_{c,c}}) \leq P_c$: In this case, λ_1 will take an arbitrarily large value and $\lambda_2 = 0$ to drive $g(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}})$ to $-\infty$.
- $\text{Tr}(\boldsymbol{\Sigma_p}) \leq P_p$ and $\text{Tr}(\boldsymbol{\Sigma_{c,p}}) + \text{Tr}(\boldsymbol{\Sigma_{c,c}}) > P_c$: In this case, $\lambda_1 = 0$ and λ_2 will take an arbitrarily large value to drive $g(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}})$ to $-\infty$.
- $\text{Tr}(\boldsymbol{\Sigma_p}) > P_p$ and $\text{Tr}(\boldsymbol{\Sigma_{c,p}}) + \text{Tr}(\boldsymbol{\Sigma_{c,c}}) > P_c$: In this case, λ_1 and λ_2 will take arbitrarily large values to drive $g(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}})$ to $-\infty$.

When both the covariance matrices satisfy the power constraints with inequality, then $\lambda_1 = \lambda_2 = 0$. This is because, $\text{Tr}(\boldsymbol{\Sigma_p}) - P_p$ and $\text{Tr}(\boldsymbol{\Sigma_{c,p}} + \boldsymbol{\Sigma_{c,c}}) - P_c$ are both negative. Hence, for any positive value of λ_1 or λ_2, $L(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \lambda_1, \lambda_2) \geq L(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, 0, 0)$.

When one of the power constraint is satisfied with equality, say $\text{Tr}(\boldsymbol{\Sigma_p}) - P_p = 0$ and the other power constraint is satisfied with inequality $\text{Tr}(\boldsymbol{\Sigma_{c,p}} + \boldsymbol{\Sigma_{c,c}}) - P_c < 0$, then, we will have $\lambda_2 = 0$ and λ_1 will be some real number. In any case, we still have $\lambda_1(\text{Tr}(\boldsymbol{\Sigma_p}) - P_p) = \lambda_2(\text{Tr}(\boldsymbol{\Sigma_{c,p}} + \boldsymbol{\Sigma_{c,c}}) - P_c) = 0$.

Similarly, when the first constraint is satisfied with inequality, and the second constraint satisfied with equality, we have $\lambda_1 = 0$ and λ_2 is some non negative real number. We have $\lambda_1(\text{Tr}(\boldsymbol{\Sigma_p}) - P_p) = \lambda_2(\text{Tr}(\boldsymbol{\Sigma_{c,p}} + \boldsymbol{\Sigma_{c,c}}) - P_c) = 0$.

Finally, if both the power constraints are satisfied with equality, λ_1 and λ_2 are some non-negative real numbers. And $\lambda_1(\text{Tr}(\boldsymbol{\Sigma_p}) - P_p) = \lambda_2(\text{Tr}(\boldsymbol{\Sigma_{c,p}} + \boldsymbol{\Sigma_{c,c}}) - P_c) = 0$.

Hence, in all the cases, the complementary slackness conditions are satisfied. Hence, the optimal solution of the optimization problem (28) satisfy the power constraints and the objective function reduces to that of optimization problem (25). Hence, both the optimization problems have the same optimal values. That is, $M = U$. ∎

Next, we find the optimum value of $\mu R_p + R_c$ over all the rate pairs that are in the region $\mathcal{R}^{\alpha}_{part,out}$ described by (12). This is done by solving the following optimization problem:

$$\sup_{((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}})} \mu R_p + R_c \qquad (29)$$
$$\text{such that} \quad ((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}}) \in \mathcal{R}^{\alpha}_{part,conv,rate}$$
$$\text{Tr}(\boldsymbol{\Sigma_{c,c}}) + \text{Tr}(\mathbf{Q_p}) \le \alpha P_c + P_p,$$

where $\mathcal{R}^{\alpha}_{part,conv,rate}$ is the set of quadruples $((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}})$ described by

$$\left\{ \begin{array}{l} ((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}}): \\ R_p \ge 0, R_c \ge 0, \mathbf{Q_p} \succeq \mathbf{0}, \boldsymbol{\Sigma_{c,c}} \succeq \mathbf{0} \\ R_p \le \log\left|\mathbf{I} + \mathbf{G}_\alpha \mathbf{Q_p} \mathbf{G}_\alpha^\dagger + \frac{1}{\alpha}\mathbf{H_{c,p}}\boldsymbol{\Sigma_{c,c}}\mathbf{H_{c,p}^\dagger}\right| \\ \qquad - \log\left|\mathbf{I} + \frac{1}{\alpha}\mathbf{H_{c,p}}\boldsymbol{\Sigma_{c,c}}\mathbf{H_{c,p}^\dagger}\right| \\ R_c \le \log\left|\mathbf{I} + \frac{1}{\alpha}\mathbf{H_{c,c}}\boldsymbol{\Sigma_{c,c}}\mathbf{H_{c,c}^\dagger}\right| \end{array} \right\}. \quad (30)$$

We let the optimal solution of (29) to be denoted by $N(\alpha)$. Let $N = \min_{\alpha>0} N(\alpha)$ and

$$\alpha^* = \arg\min_{\alpha>0} N(\alpha). \qquad (31)$$

We show in Lemma 6.2 that $\alpha^* \in (0,\infty)$ exists. Then, N is given by the optimum value of the following $\inf\sup$ optimization problem

$$\inf_{\alpha>0} \quad \sup_{((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}})} \mu R_p + R_c \qquad (32)$$
$$\text{such that} \quad ((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}}) \in \mathcal{R}^{\alpha}_{part,conv,rate}$$
$$\text{Tr}(\boldsymbol{\Sigma_{c,c}}) + \text{Tr}(\mathbf{Q_p}) \le \alpha P_c + P_p.$$

The infimum constraint $\alpha > 0$ is not a compact set. We modify the constraint on α to $\alpha \in \mathbb{R}^+ \cup \{0,\infty\}$. This is done to compactify the set by adding two extra symbols 0 and ∞. The point zero is added to make the set closed. The process of adding the point ∞ is called one point compactification. Details on one point compactification can be found in [24,

Section 2.8]. The new space $\alpha \in \mathbb{R}^+ \cup \{0,\infty\}$ is compact and Hausdorff.

The optimization problem after changing the constraint set on α becomes

$$N_1 = \inf_{\alpha \in \mathbb{R}^+ \cup \{0,\infty\}} \quad \sup_{((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}})} \mu R_p + R_c \qquad (33)$$
$$\text{such that} \quad ((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}}) \in \mathcal{R}^{\alpha}_{part,conv,rate}$$
$$\text{Tr}(\boldsymbol{\Sigma_{c,c}}) + \text{Tr}(\mathbf{Q_p}) \le \alpha P_c + P_p.$$

We show that adding the two points 0 and ∞ to the constraint set on α does not change the optimum value of the optimization problem. This result is formally stated and proved in the following lemma.

Lemma 6.2: The optimum value of the optimization problem given by (32), N is equal to the optimum value of the optimization problem described by (33), N_1. That is, $N = N_1$.

Proof : For any $\alpha \in \mathbb{R}^+ \cup \{0,\infty\}$, we let $h(\alpha)$ to denote the value of the inner $\sup$ problem. That is,

$$h(\alpha) = \sup_{((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}})} \mu R_p + R_c \qquad (34)$$
$$\text{such that} \quad ((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}}) \in \mathcal{R}^{\alpha}_{part,conv,rate}$$
$$\text{Tr}(\boldsymbol{\Sigma_{c,c}}) + \text{Tr}(\mathbf{Q_p}) \le P_p + \alpha P_c.$$

We show that $\liminf_{\alpha\to 0} h(\alpha) = \liminf_{\alpha\to\infty} h(\alpha) = \infty$.

Letting $\alpha \to 0$, we put all the power in $\boldsymbol{\Sigma_{c,c}}$. That is, we choose $\boldsymbol{\Sigma_p} = \mathbf{0}$, $\boldsymbol{\Sigma_{c,p}} = \mathbf{0}$, $\mathbf{Q} = \mathbf{0}$ and $\boldsymbol{\Sigma_{c,c}} = \frac{P_p+\alpha P_c}{n_{c,t}}\mathbf{I}_{n_{c,t}}$. Also, we take

$$R_p = 0 \text{ and } R_c = \log\left|\mathbf{I} + \frac{1}{\alpha}\frac{P_p+\alpha P_c}{n_{c,t}}\mathbf{H_{c,c}}\mathbf{H_{c,c}^\dagger}\right|.$$

It follows from (30) that $((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}}) \in \mathcal{R}^{\alpha}_{part,conv,rate}$. Also, $\text{Tr}(\mathbf{Q_p}) + \text{Tr}(\boldsymbol{\Sigma_{c,c}}) = P_p + \alpha P_c$. Hence, $((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}})$ satisfy all the necessary constraints of (34). Substituting these particular values of $((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}})$, we get a lower bound on $h(\alpha)$. That is,

$$\liminf_{\alpha\to 0} h(\alpha) \ge \liminf_{\alpha\to 0} \log\left|\mathbf{I} + \frac{1}{\alpha}\frac{P_p+\alpha P_c}{n_{c,t}}\mathbf{H_{c,c}}\mathbf{H_{c,c}^\dagger}\right|$$
$$= \infty. \qquad (35)$$

Next, we look at the situation when $\alpha \to \infty$. In this case, we put all the power in $\boldsymbol{\Sigma_p}$. That is, we choose $\boldsymbol{\Sigma_p} = \frac{P_p+\alpha P_c}{n_{p,t}}\mathbf{I}_{n_{p,t}}$, $\boldsymbol{\Sigma_{c,p}} = \mathbf{0}$, $\boldsymbol{\Sigma_{c,c}} = \mathbf{0}$ and $\mathbf{Q} = \mathbf{0}$. We also choose

$$R_c = 0 \text{ and } R_p = \log\left|\mathbf{I} + \frac{P_p+\alpha P_c}{n_{p,t}}\mathbf{H_{p,p}}\mathbf{H_{p,p}^\dagger}\right|.$$

These values of $((R_p,R_c),\mathbf{Q_p},\boldsymbol{\Sigma_{c,c}})$ satisfy all the necessary constraints of (34). Hence, we have

$$\liminf_{\alpha\to\infty} h(\alpha) \ge \liminf_{\alpha\to\infty} \mu \log\left|\mathbf{I} + \frac{P_p+\alpha P_c}{n_{p,t}}\mathbf{H_{p,p}}\mathbf{H_{p,p}^\dagger}\right|$$
$$= \infty. \qquad (36)$$

Hence, $h(\alpha) = \infty$ when $\alpha = 0$ or $\alpha = \infty$. Also, when $\alpha \in \mathbb{R}^+$, $h(\alpha) < \infty$. Hence, the optimum value of (33) is reached when α is neither 0 nor ∞. Hence, $N = N_1$. ∎

As $\mathbf{Q_p}$ is the covariance matrix of the codeword $\mathbf{X}(i), i = 1, \ldots, n$ for the primary user, it can be written as

$$\mathbf{Q_p} = \begin{pmatrix} \mathbf{\Sigma_p} & \mathbf{Q} \\ \mathbf{Q}^\dagger & \mathbf{\Sigma_{c,p}} \end{pmatrix}. \tag{37}$$

It is easy to see that the set $R_{part,conv}^\alpha$ described in (11) can also be written as

$$\left\{ \begin{array}{l} \left((R_p, R_c), \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{Q}, \mathbf{\Sigma_{c,c}} \right) : \\ R_p \geq 0, R_c \geq 0, \mathbf{\Sigma_p} \succeq \mathbf{0}, \mathbf{\Sigma_{c,p}} \succeq \mathbf{0}, \mathbf{\Sigma_{c,c}} \succeq \mathbf{0} \\ R_p \leq \log \left| \mathbf{I} + \mathbf{G}\mathbf{Q_p}\mathbf{G}^\dagger + \mathbf{H_{c,p}}\mathbf{\Sigma_{c,c}}\mathbf{H_{c,p}}^\dagger \right| \\ \qquad - \log \left| \mathbf{I} + \mathbf{H_{c,p}}\mathbf{\Sigma_{c,c}}\mathbf{H_{c,p}}^\dagger \right| \\ R_c \leq \log \left| \mathbf{I} + \mathbf{H_{c,c}}\mathbf{\Sigma_{c,c}}\mathbf{H_{c,c}}^\dagger \right| \\ \mathrm{Tr}(\mathbf{\Sigma_p}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,p}}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,c}}) \leq P_p + \alpha P_c \end{array} \right\}. \tag{38}$$

where $\mathbf{G} = [\mathbf{H_{p,p}} \ \mathbf{H_{c,p}}]$. This is done by transforming $\mathbf{Q}, \mathbf{\Sigma_{c,p}}, \mathbf{\Sigma_{c,c}}$ into $\sqrt{\alpha}\mathbf{Q}, \alpha\mathbf{\Sigma_{c,p}}, \alpha\mathbf{\Sigma_{c,c}}$ respectively. We define $\mathcal{R}_{part,conv,rate}$ as the set described by

$$\left\{ \begin{array}{l} ((R_p, R_c), \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{Q}, \mathbf{\Sigma_{c,c}}) : \\ R_p \geq 0, R_c \geq 0, \mathbf{\Sigma_p} \succeq \mathbf{0}, \mathbf{\Sigma_{c,p}} \succeq \mathbf{0}, \mathbf{\Sigma_{c,c}} \succeq \mathbf{0} \\ R_p \leq \log \left| \mathbf{I} + \mathbf{G}\mathbf{Q_p}\mathbf{G}^\dagger + \mathbf{H_{c,p}}\mathbf{\Sigma_{c,c}}\mathbf{H_{c,p}}^\dagger \right| \\ \qquad - \log \left| \mathbf{I} + \mathbf{H_{c,p}}\mathbf{\Sigma_{c,c}}\mathbf{H_{c,p}}^\dagger \right| \\ R_c \leq \log \left| \mathbf{I} + \mathbf{H_{c,c}}\mathbf{\Sigma_{c,c}}\mathbf{H_{c,c}}^\dagger \right|, \\ \mathbf{Q_p} = \begin{pmatrix} \mathbf{\Sigma_p} & \mathbf{Q} \\ \mathbf{Q}^\dagger & \mathbf{\Sigma_{c,p}} \end{pmatrix} \end{array} \right\}. \tag{39}$$

Hence, the optimization problem (33) can be written as

$$N = \inf_{\alpha \in \mathbb{R}^+ \cup \{0,\infty\}} \sup_{((R_p, R_c), \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{Q}, \mathbf{\Sigma_{c,c}})} \mu R_p + R_c \tag{40}$$

such that $((R_p, R_c), \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{Q}, \mathbf{\Sigma_{c,c}}) \in \mathcal{R}_{part,conv,rate}$
$$\mathrm{Tr}(\mathbf{\Sigma_p}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,p}}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,c}}) \leq P_p + \alpha P_c.$$

We state the following lemma for switching min and max in minimax problems. The lemma is described and proved in Theorem 2 in [25].

Lemma 6.3: (Ky-Fan's minimax switching theorem [25, Thm. 2]) Let X be a compact Hausdorff space and Y an arbitrary set (not topologized). Let f be a real-valued function on $X \times Y$ such that, for every $y \in Y$, $f(x, y)$ is lower semi continuous on X. If f is convex on X and concave on Y, then

$$\inf_{x \in X} \sup_{y \in Y} f(x, y) = \sup_{y \in Y} \inf_{x \in X} f(x, y).^2 \tag{41}$$

We see that the objective function $\mu R_p + R_c$ is concave with respect to the maximizing variables $((R_p, R_c, \mathbf{Q_p}, \mathbf{\Sigma_{c,c}})$ and convex with respect to the minimizing variable α. The

---

^{2}In (49), the inf can be replaced with min, but we use inf throughout to maintain continuity and to avoid confusion.

constraint space $\alpha \in \mathbb{R}^+ \cup \{0, \infty\}$ is compact and Hausdorff [24, Section 2.8].

Hence, all the conditions of the lemma are satisfied. Hence, by Ky-Fan's mini-max switching theorem [25], we can interchange the sup and inf without affecting the optimum value. Hence,

$$N = \sup_{((R_p, R_c), \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{Q}, \mathbf{\Sigma_{c,c}})} \inf_{\alpha \in \mathbb{R}^+ \cup \{0,\infty\}} \mu R_p + R_c \tag{42}$$

such that $((R_p, R_c), \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{Q}, \mathbf{\Sigma_{c,c}}) \in \mathcal{R}_{part,conv,rate}$
$$\mathrm{Tr}(\mathbf{\Sigma_p}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,p}}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,c}}) \leq P_p + \alpha P_c.$$

Similar to the functions L and g defined in (26) and (27), we define the functions $L_1(R_p, R_c, \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{\Sigma_{c,c}}, \lambda, \alpha)$ and $g_1(R_p, R_c, \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{\Sigma_{c,c}}, \alpha)$ as follows

$$L_1(R_p, R_c, \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{\Sigma_{c,c}}, \lambda, \alpha) = \mu R_p + R_c - \\ \lambda \left(\mathrm{Tr}(\mathbf{\Sigma_p}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,p}}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,c}}) - P_p - \alpha P_c \right), \tag{43}$$

$$g_1(R_p, R_c, \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{\Sigma_{c,c}}, \alpha) = \\ \inf_{\lambda \geq 0} L_1(R_p, R_c, \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{\Sigma_{c,c}}, \lambda, \alpha). \tag{44}$$

We define the following optimization problem

$$V = \sup_{(R_p, R_c, \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{Q}, \mathbf{\Sigma_{c,c}})} \inf_\alpha g_1(R_p, R_c, \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{\Sigma_{c,c}}, \alpha) \tag{45}$$

such that $((R_p, R_c), \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{Q}, \mathbf{\Sigma_{c,c}}) \in \mathcal{R}_{part,conv,rate}$
$$\alpha \in \mathbb{R}^+ \cup \{0, \infty\}.$$

Lemma 6.4: The optimum value of optimization problem (42), N is equal to the optimum value of the optimization problem (45), V.

Proof : The proof of the lemma is along the same lines as the proof of Lemma 6.1. We show that for any set of covariance matrices $\mathbf{\Sigma_p}$, $\mathbf{\Sigma_{c,p}}$ and $\mathbf{\Sigma_{c,c}}$ that do not satisfy the power constraint $\mathrm{Tr}(\mathbf{\Sigma_p}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,p}}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,c}}) \leq P_p + \alpha P_c$, $g_1(R_p, R_c, \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{\Sigma_{c,c}}, \alpha) = -\infty$. This is because, $\mathrm{Tr}(\mathbf{\Sigma_p}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,p}}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,c}}) - P_p - \alpha P_c$ is positive, and hence, λ will take an arbitrarily high value to drive $g_1(R_p, R_c, \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{\Sigma_{c,c}}, \alpha)$ to $-\infty$. Hence, the outer supremization problem will ensure that the power constraint is satisfied.

Moreover, when the power constraints are satisfied with inequality, then $\mathrm{Tr}(\mathbf{\Sigma_p}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,p}}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,c}}) - P_p - \alpha P_c$ is negative. Therefore, for any $\lambda > 0$, we have $L_1(R_p, R_c, \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{\Sigma_{c,c}}, \lambda, \alpha) > L_1(R_p, R_c, \mathbf{\Sigma_p}, \mathbf{\Sigma_{c,p}}, \mathbf{\Sigma_{c,c}}, 0, \alpha)$. Hence, λ will take the value zero. When the power constraint is satisfied with equality, then $\mathrm{Tr}(\mathbf{\Sigma_p}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,p}}) + \alpha \mathrm{Tr}(\mathbf{\Sigma_{c,c}}) - P_p - \alpha P_c = 0$. Then, λ will take some non negative real number. Hence, the complementary slackness condition is satisfied. Hence, the optimal solution of the optimization problem satisfy the power constraint and the objective function reduces to that of

(42). It follows that, the optimum value of the optimization problem (42), N is the same as the optimum value of the optimization problem (45), V.

Next, we show that the optimum value of the optimization problem (28), U is an upper bound on the optimal value of the optimization problem (45), V.

Lemma 6.5: The optimal value of (28), U is an upper bound on the optimal value of (42), V.

Proof: Both the optimization problems are $\sup \min$ problems. For any $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$, we can choose $\lambda = \lambda_1$ and $\alpha = \lambda_2/\lambda_1$ so that $L_1(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \lambda, \alpha) = L(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \lambda_1, \lambda_2)$. Hence, for any $((R_p, R_c), \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}})$,

$$\inf_{\lambda \geq 0, \alpha \in \mathbb{R}^+ \cup \{0, \infty\}} L_1(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \lambda, \alpha) \leq$$
$$\inf_{\lambda_1 \geq 0, \lambda_2 \geq 0} L(R_p, R_c, \boldsymbol{\Sigma_p}, \boldsymbol{\Sigma_{c,p}}, \boldsymbol{\Sigma_{c,c}}, \lambda_1, \lambda_2). \quad (46)$$

Also, $\mathcal{R}_{part,conv,rate} = \mathcal{R}_{ach,rate}$. Hence, it follows that $V \leq U$.   ∎

We can now prove Theorem 3.4.

Proof of Theorem 3.4 : Let $\mu \geq 1$. The proof of the theorem follows directly from Lemmas 6.1, 6.4 and 6.5. From Lemma 6.1, we have that the optimum value of the optimization problem (25), M equals the optimum value of optimization problem (28), U. From Lemma 6.4, we have that the optimum value of optimization problem (42), N equals the optimum value of the optimization problem (45), V. M is the solution of the optimum $\mu R_p + R_c$ over the achievable region and N is the solution of the optimum $\mu R_p + R_c$ over $\mathcal{R}_{part,out}^{\alpha}$ described in (12). Hence if the condition given by (13) is satisfied for α^* given by (31), $M \leq N$. From Lemma 6.5, we also have $V \leq U$. Hence, we have that the optimal value of the original optimization problem (25), M is equal to the optimal value of the optimization problem described by (42), N. Hence, the achievable region $\mathcal{R}_{in}$ is μ-sum optimal.   ∎

VII. NUMERICAL RESULTS

In this section, we provide some numerical results on the capacity region of the MIMO cognitive channel. We consider a MIMO cognitive system where the licensed and cognitive transmitters have one antenna each, and the licensed and cognitive receivers have one and two antennas respectively. We assume that the channel coefficients are real and also restrict ourself to real inputs and outputs. We generate the channel values randomly

$$\mathbf{H_{p,p}} = 1.4435, \quad \mathbf{H_{p,c}} = \begin{bmatrix} -0.3510 \\ 0.6232 \end{bmatrix},$$

$$\mathbf{H_{c,p}} = 0.799, \quad \mathbf{H_{c,c}} = \begin{bmatrix} 0.9409 \\ -0.9921 \end{bmatrix}.$$

We assume a power constraint of 5 at the licensed and cognitive transmitters. In Figure 8, we plot the achievable region, $\mathcal{R}_{in}$ and the region $\mathcal{R}_{part,out}^{\alpha}$ for different values of α.
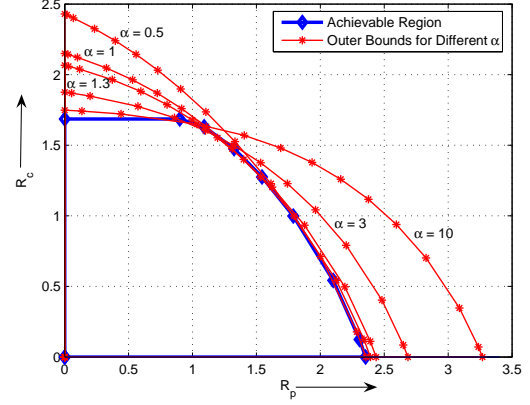

Fig. 8. Plot of Achievable Region $\mathcal{R}_{in}$ and partial outer bounds $\mathcal{R}_{part,out}^{\alpha}$ for different values of α.

Figure 8 shows how $\mathcal{R}_{part,out}^{\alpha}$ intersects with $\mathcal{R}_{in}$ at different points for different values of α.

Next, we find the maximum value of rate than can be supported by the licensed user in the example we considered. In both the achievable region and the outer bound, this corresponds to maximizing the μ-sum $\mu R_p + R_c$ when $\mu \to \infty$. This would correspond to using all the power to support the licensed user. Note that the maximum value of R_p in the set described by $\mathcal{R}_{part,out}^{\alpha}$ is an upper bound on the maximum value of R_p in the set $\mathcal{R}_{in}$ for all values of $\alpha > 0$, irrespective of the channel parameters.

Maximizing R_p over $\mathcal{R}_{in}$: The cognitive transmitter uses all its power for helping the licensed user. That is $\text{Tr}(\boldsymbol{\Sigma_{c,p}}) = P_c$. This then reduces to a MIMO channel with channel matrix given by $\mathbf{G} = \begin{bmatrix} \mathbf{H_{p,p}} & \mathbf{H_{c,p}} \end{bmatrix}$. The licensed transmitter has a power constraint of P_p and the cognitive transmitter has a power constraint of P_c. Applying this to our example channel, we have $\mathbf{G} = \begin{bmatrix} 1.4435 & 0.799 \end{bmatrix}$. The optimum covariance matrix is of the form

$$\boldsymbol{\Sigma_{p,net}} = \begin{bmatrix} 5 & 5\rho \\ 5\rho & 5 \end{bmatrix},$$

where ρ is the correlation between the two transmitters. Therefore, the rate achieved by the licensed user is

$$R_p(\rho) = \frac{1}{2} \log(1 + \mathbf{G}\boldsymbol{\Sigma_{p,net}}\mathbf{G}^{\dagger}).$$

The maximum rate is attained at $\rho = 1$ and the maximum value of R_p is 2.3542.

Maximizing R_p over $\mathcal{R}_{part,out}^{\alpha}$: For a given α, this reduces to a single user MIMO channel with $\mathbf{G}_{\alpha} = \begin{bmatrix} \mathbf{H_{p,p}} & \mathbf{H_{c,p}}/\sqrt{\alpha} \end{bmatrix}$ and a sum power constraint of $P_p + \alpha P_c$. Note that, there is a significant difference between the two single user MIMO channels. The MIMO channel that we considered when solving the maximum value of R_p in the achievable region had individual power constraints at the licensed and cognitive transmitters. However, the MIMO channel we obtain when solving for the maximum value of R_p over $\mathcal{R}_{part,out}^{\alpha}$ has a sum power constraint. This is

a conventional MIMO channel and the optimum covariance matrix is obtained by water-filling. For a given α, the best R_p is got by

$$\max R_p(\alpha) = \frac{1}{2} \log |\mathbf{I} + \mathbf{G}_\alpha \mathbf{\Sigma}_{\mathbf{p,net}} \mathbf{G}_\alpha|$$
$$\text{such that } \text{Tr}(\mathbf{\Sigma}_{\mathbf{p,net}}) \leq P_p + \alpha P_c.$$

It is easy to solve this problem if we look at the flipped channel $\mathbf{G}_\alpha^\dagger$. The capacity of the flipped channel is given by

$$R_p(\alpha) = \frac{1}{2} \log \left| \mathbf{I} + \mathbf{G}_\alpha^\dagger (P_p + \alpha P_c) \mathbf{G}_\alpha \right|$$
$$= \frac{1}{2} \log \left(1 + (P_p + \alpha P_c) \mathbf{G}_\alpha \mathbf{G}_\alpha^\dagger \right).$$

Note that $R_p(\alpha)$ is an outer bound on the maximum value of R_p. The best upper bound is got by minimizing over all possible values of α. The optimum value of α is got by solving a cubic equation $2(0.799)^2 \alpha^3 + (0.799)^2 \alpha^2 - 1.4435^2 = 0$, and its approximate value is 0.9689.

VIII. CONCLUSIONS

In this paper, we derived an achievable region, $\mathcal{R}_{in}$ given by (4) and an outer bound, $\mathcal{R}_{out}^{\alpha, \mathbf{\Sigma_z}}$ given by (8) for the MIMO cognitive channel. We describe conditions when the achievable region is μ-sum optimal for any $\mu \geq 1$. In particular, for any $\mu \geq 1$, there exists $\alpha^* \in (0, \infty)$, such that if the region given by $\mathcal{R}_{part,out}^{\alpha^*}$ optimizes the $\mu-$ sum rate of the SMBC (for that particular α^*), then the achievable region achieves the μ-sum capacity of the MCC.

REFERENCES

[1] S. Sridharan and S. Vishwanath, "On the capacity of a class of MIMO cognitive radios," in *Proc. Information Theory Workshop (ITW) 2007, Lake Tahoe*, Sept. 2007, pp. 384–389.

[2] *Federal Communications Commision, Cognitive Radio Technologies Proceeding (CRTP), ET Docket No. 03-108.* http://www.fcc.gov/oet/cognitiveradio/.

[3] J. Mitola, *Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio*. PhD Dissertation, KTH, Stockholm, Sweeden, December 2000.

[4] A. Ghasemi and E. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proc. First IEEE International Symp. on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, 8-11 Nov. 2005, pp. 131–136.

[5] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on selected areas in Communications*, vol. 23, no. 2, pp. 201– 220, Feb 2005.

[6] S. Srinivasa, S. Jafar, and N. Jindal, "On the capacity of the cognitive tracking channel," in *IEEE International Symposium on Information Theory*, July 2006, pp. 2077–2081.

[7] S. A. Jafar and S. Srinivasa, "Capacity limits of cognitive radio with distributed and dynamic spectral activity," *IEEE Journal on selected Areas in Communications*, vol. 25, no. 3, pp. 529–537, April 2007.

[8] S. Srinivasa and S. A. Jafar, "The throughput potential of cognitive radio: A theoretical perspective," *IEEE Communications Magazine*, vol. 45, no. 5, 2007.

[9] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[10] A. Carleial, "Interference channels," *IEEE Trans. Inform. Theory*, vol. 24, no. 1, pp. 60–70, Jan. 1978.

[11] M. Costa, "On the Gaussian interference channel," *IEEE Trans. Inform. Theory*, vol. 31, no. 5, pp. 607–615, Sep 1985.

[12] H. Sato, "The capacity of the Gaussian interference channel under strong interference (corresp.)," *IEEE Trans. Inform. Theory*, vol. 27, no. 6, pp. 786–788, Nov 1981.

[13] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inform. Theory*, vol. 27, no. 1, pp. 49–60, Jan 1981.

[14] I. Maric, R. Yates, and G. Kramer, "The capacity region of the strong interference channel with common information," in *Proc. Thirty-Ninth Asilomar Conference on Signals, Systems and Computers*, October 28 - November 1, 2005, pp. 1737–1741.

[15] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," submitted to IEEE Trans. Inform. Theory. Preprint available at http://www.ifp.uiuc.edu~pramodv/pubs.html.

[16] W. Wu, S. Vishwanath, and A. Arapostathis, "On the capacity of Gaussian weak interference channels with degraded message sets," accepted at IEEE Trans. Inform. Theory. Preprint available at http://users.ece.utexas.edu~wwu/publications.html.

[17] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3936–3964, Sept. 2006.

[18] M. Mohseni and J. M. Cioffi, "A proof of the converse for the capacity of Gaussian MIMO broadcast channels," in *Proc. IEEE Intl. Symp. on Inform. Theory, Seattle, 2006*, July 2006, pp. 881–885.

[19] H. Weingarten, T. Liu, S. S. (Shitz), Y. Steinberg, and P. Viswanath, "The capacity region of the degraded multiple input multiple output compound broadcast channel," submitted to IEEE Trans. Inform. Theory. Preprint available at http://www.ece.tamu.edu~tieliu/pub/dc07.pdf.

[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley Series in Telecommunications, July 2006.

[21] M. Costa, "Writing on dirty paper (corresp.)," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[22] W. Yu and J. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1875–1892, Sept. 2004.

[23] S. Vishwanath, G. Kramer, S. Shamai, S. Jafar, and A. Goldsmith, "Capacity bounds for Gaussian vector broadcast channels," in *DIMACS Workshop on Signal Processing for Wireless Transmission, Rutgers, NJ*, October 2002.

[24] R. M. Dudley, *Real analysis and probability*. Cambridge University Press, 2002.

[25] K. Fan, "Minimax theorems," in *Proc. Nat. Acad. Sci, 39*, 1953, pp. 42–47.